Algorithm Theoretical Basis Document

D1 The Next Generation Model

Contents

List of Acronyms

CCM: Community Climate Model

CGD: Climate and Global Dynamics Division

DAO: Data Assimilation Office

GCM: General Circulation Model

GRIP: GCM reality Inter-comparison Project

LLNL: Lawrence Livermore National Laboratory

LSM: Land Surface Model

MPI: Message Passing Interface

MPP: Massive Parallel computing Platform

NCAR: National Center for Atmospheric Research

NCEP: National Center for Environment Prediction

NWP: Numerical Weather Prediction

1. Introduction

As a long-term strategic effort to sustain a state-of-the-art of the modeling system, the Data Assimilation Office (DAO) is collaborating with NCAR Climate and Global Dynamics Division (CGD) in the joint development of a global model. The ultimate goal of this collaboration is to build a unified climate, numerical weather prediction (NWP), and chemistry-transport model, which is suitable for global data assimilation of the physical and chemical state of the Earth's atmosphere. To this end, a high standard must be set for the integrity of the chosen scientific algorithms, and a rigorous software engineering practice must be enforced for the implementation.

The DAO's major contribution to the joint model is in the development of a next-generation dynamical core for global models capable of resolving atmospheric motions from meso- to planetary- scale with a high throughput on multiple-processor, distributed-memory computing platforms. The basic algorithms are derived and evolved from the modern high-resolution finite-volume algorithms pioneered by Van Leer (1977) and Colella and Woodward (1984), which are one-dimensional algorithms designed primarily for astrophysical and aerospace engineering applications requiring the resolution of sharp gradients (e.g., shock waves).

While the atmospheric modeling community is largely aware of the advantages of finite-volume algorithms in chemistry-transport applications (reviewed by Rood 1987), it was not until the works of Lin and Rood [as documented in a series of papers and conference and workshop proceedings, e.g., Lin and Rood 1996 (LR96), and Lin and Rood 1997 (LR97), Lin 1997 (L97), Lin and Rood 1998 (LR98), and Lin and Rood 1999 (LR99)] that the finite-volume based algorithms became viewed as a viable candidate to replace the traditional finite-difference methods, the spectral method, and the widely used advective-form semi-Lagrangian methods (*e.g.*, Staniforth and Cote 1991) for general circulation models. LR96 extended the mass conserving and monotonicity preserving (*i.e.*, Gibbs' oscillation free) finite-volume transport algorithms to multi-dimensions and "semi-Lagrangian" to avoid the "Pole Courant-number Problem." LR96 focused on the tracer transport problem on the sphere, but with these extensions the viability of the approach for global geophysical glows became real. LR97 subsequently developed an explicit algorithm for PV (Potential Vorticity) transport for the shallow water dynamical framework via a computationally efficient "reversed engineering approach" (see LR97 for

details). This development provided an advection scheme capable of not only high accuracy, but also assured physical consistency, at least in a 2D shallow water flow, of the conservation equations that govern atmospheric dynamics.

The extension of the finite-volume algorithms to the 3D primitive equations followed from the development of an accurate finite-volume integration method for computing the pressure gradients in general vertical coordinates (L97). This method for computing the pressure gradient is combined with a vertical Lagrangian control-volume discretization (LR98 and LR99) that eliminates the need to explicitly model the vertical transport process. This unique model configuration also opens up more favorable attributes for parallel computing. The development of the finite-volume dynamical core will be described in detail in section 2. The Lagrangian control-volume vertical coordinate and the associated mass, momentum, and total energy conservative mapping algorithm will be described in section 3.

NCAR CGD's contributions to the joint model development are in the physical parameterizations and a coupled land-surface model (LSM) component (Bonan 1996). The physical parameterizations of the joint model are based on the Community Climate Model (CCM3), a state-of-the-art atmospheric model with a broad user base in both the university and national laboratory communities. The performance of CCM3 is documented in a special issue of *the Journal of Climate* (Vol. 11, no. 6, June 1998), which details both strengths and weaknesses in the model. CCM3 is presently based on a spectral dynamical core. DAO's finite-volume dynamical core is being considered a candidate for the inclusion into NCAR CCM4, since it addresses a range of deficiencies in the NCAR CCM spectral numerics that are becoming more consequential as the physical parameterizations become more interactive. For example, the spectral ringing west of the Andes in CCM3 is clearly diagnosed to modulate the clouds and surface flux exchanges and may partially contribute to the lack of marine stratus. The ringing is absent when the new core is used. Further, the conservation characteristics of the finite-volume dynamical core also increase the integrity of the simulation of trace species, alleviating the impact that the mass fixer in the current CCM has on some chemistry problems.

A successful joint NASA/NCAR model would be a significant step toward the development of a multi-agency modeling capability, and could point the way for the

inclusion of other interested organizations (e.g., NAVY/NRL and various laboratories in NOAA and DOE). Therefore, an ambitious long term goal of this joint model development would be the appropriate unification/consolidation of various global models previously developed independently for climate modeling, climate prediction, weather prediction, and data assimilation at different institutions within the U.S. DAO's focus will be in the model's applications to data assimilation for climate and chemistry applications, while NCAR's focus will be on the model's climate simulations. The validation and applications of the model by the two groups should help to more thoroughly integrate the use of observations in climate model development. In as much as data assimilation relies on weather forecasting, we anticipate that the activity will also help to reconcile the sometimes-conflicting problems of weather forecasting and climate simulation. This part of the ATBD outlines the fundamentals of the joint model with special focus on the scientific attributes of the finite-volume dynamical core. Some exemplary results from the joint model will be presented in section 6. The next part describes the prototype assimilation system using the new model.

2. The development of the finite-volume dynamical core

The development of the finite-volume algorithms for global models at NASA/GSFC started in the late 80s and early 90s with focus on transport process of chemistry constituents (*e.g.*, Rood 1987; Rood et al. 1991; Allen et al. 1991) and water vapor (Lin et al. 1994). The quality and physical integrity of constituent simulations benefited so greatly from the application of modern, physically-based, flux-form transport algorithms that the field was able to move away from space-time averaged model-data inter-comparisons to event-by-event evaluation of in situ constituent observations (*e.g.*, Rood et al. 1991; Allen et al. 1996; Douglass et al., 1997, and references therein). The conservation equations for atmospheric motion should, similarly, benefit from an increased level of physical rigor in their numerical expression, allowing more quantitative evaluation of the terms commonly referred to as "physics" in general circulation models. This premise motivated us to undertake the challenge of developing a computationally efficient, physically-based finite-volume dynamical core for global models, which is considerably more complicated than the constituent transport problems we considered and successfully modeled previously.

The development started with the existing foundation of finite-volume schemes designed originally for entirely different applications (e.g., Woodward and Colella 1984). As stated in the Introduction, most of the finite-volume schemes are one dimensional by construction and developed to perform well in sound wave dominated flows (e.g., shock waves), which have essentially no meteorological significance on the global scale. Furthermore, the large-scale atmospheric motions are basically hydrostatic and nearly incompressible. As such, the standard "Riemann solver" (*e.g.*, Carpenter et al. 1991) would not be efficient or applicable. Further, the directional splitting necessary for using the 1D finite-volume algorithms would produce unacceptably large errors, particularly near the poles where the splitting errors would be amplified by the convergence of the meridians (for the effects of the operator splitting near the poles, see Allen et al. 1991 and Lin et al. 1994).

The primitive equations governing geophysical flows represent basic conservation laws. A standard way to derive these laws from first principles is the consideration of finite control volumes (as can be found in most textbooks on fluid dynamics). It was an early design decision that the physical conservation laws be built, as much as possible, directly into the discrete finite-volumes. Therefore, the standard procedure of approximating the differential equations with mathematically motivated techniques (e.g., finite difference, finite element, and spectral methods) is bypassed in favor of piecewise-continuous integral representations of control volumes. A good example of the process is the derivation of the finite-volume integration method for computing the pressure gradient for general terrain following (Lin, 1997). The standard mathematical transformation (via chain rule) of the pressure gradient forcing in terrain following coordinates results in two large-in-magnitude terms with opposite sign. A straightforward application of numerical techniques (e.g., center differencing) to these two terms will not only produce large errors, but also results in violation of a basic physical law because of a spurious numerical component of the pressure balance. The finite-volume integration scheme of Lin (1997) avoided the mathematical transformation by integrating around the arbitrarily shaped finite-volume to accurately determine the pressure gradient forcing that is physically consistent for the finite volume under consideration. Compared with traditional approaches, spurious forces associated with topography were significantly reduced.

Central to the accuracy and computational efficiency of the finite-volume algorithms is the number of degrees of freedom that prescribe the implied subgrid distribution. A continuous representation of the subgrid distribution is desired to enable the analytic integration of the basic conservation laws within the volume (a volume is equivalent to a single grid point in center differencing schemes). The first order upwind scheme, for example, has zero degree of freedom as it is assumed that the subgrid distribution is piecewise constant having the same value as the given volume-mean. The second order finite-volume scheme assumes a piecewise linear subgrid distribution, which allows an extra degree of freedom for the specification of the "slope" of the linear distribution to improve the accuracy of the volume integration (see, for example, Lin et al. 1994). The Piece-wise Parabolic Method (PPM, Colella and Woodward 1984) has two extra degrees of freedom in the construction of the second order polynomial within the volume, and as a result, the accuracy is significantly enhanced. Note that the subgrid distributions are compact, and applied only within the volume. The accuracy is, therefore, significantly better than the order of the chosen polynomials. Because the PPM strikes an good balance between computational efficiency and accuracy, it is the basic 1D foundation. The scheme has most of the desirable physical attributes (e.g. mass conserving and monotonicity preserving) in 1D; however, it is difficult to maintain these attributes in the presence of the directional splitting methods used to extend these schemes to the multi-dimensional problem of modeling the dynamics of the Earth's atmosphere.

The extension of the scheme to 3D on a sphere is presented in LR96 and is called the “Flux-Form Semi-Lagrangian” (FFSL) scheme. Briefly, two antisymmetric methods used to compute the advective fluxes in the meridional and zonal directions are averaged to yield a symmetric method. The operators that are the cross terms in the resulting symmetric method are then represented by their advective form. This yields a scheme that is conservative, consistent (i.e. background fluid mass is conserved in the constituent problem), and stable. Equally important, the so-called "Pole-Courant number problem" is solved, with the consideration of the contribution to fluxes from upstream volumes as far away as Courant number indicated. The resulting multi-dimensional scheme is oscillation free, mass conserving, and stable for Courant number greater than unity. While there are a large number of operations at each time step, the combined attributes of the scheme have proven to be computationally very competitive. This particular algorithm, or similarly derived schemes, are used in many chemistry-transport applications.

The LR96 paper only addressed the multi-dimensional tracer transport problem. The successful application of the FFSL algorithm to the shallow water dynamical framework is fully documented in Lin and Rood 1997 (LR97). A reversed engineering approach was devised in LR97 to achieve the design goal of consistent transport of the mass (layer thickness) and the absolute vorticity, and hence, the potential vorticity. The "reversed engineering approach" is a two-grid (C and D grids), two-step procedure to achieve the consistent transport of absolute vorticity and mass without the computational expense of using, explicitly, the Z grid (Randall 1994). Z grid methods generally require an elliptic solver. To avoid an elliptic solver, the momentum equations are formulated in the vector-invariant form to facilitate the use of the FFSL-scheme's computed absolute vorticity fluxes in the meridional and zonal directions as the forcing terms for the zonal momentum and meridional momentum, respectively (see Eq. 16 and 17 in LR97). This approach is consistent with the Z grid in that the rotational part of the flow is computed on the D grid whereas the divergent part of the flow is computed on the C grid.

The time integration scheme for treating the gravity waves on both grids is explicit and consistent with the forward-in-time nature of the FFSL transport algorithm (see LR97 for details). The allowable size of the time step, for example, for a T42-like resolution (~2.8^{o}) is 600 seconds, which is nearly half of what can be used by the semi-implicit Eulerian spectral model. This relatively long time step made the explicit shallow water algorithm computationally competitive with the traditional spectral and finite difference methods. In a 3D framework, the computational competitiveness is substantially increased because of the efficiency in parallelization, the large-time-step tracer transport, and the introduction of the Lagrangian control-volume coordinate (section 3).

The finite-volume dynamical core at the time of writing L97 utilized a sigma vertical coordinate, which requires a 3D transport algorithm. Applying the methodology of LR96, a fully 3D FFSL transport algorithm would require 6 permutations, instead of 2 as in 2D, of 1D operators to make the algorithm directionally symmetric. To reduce computational cost, a drastic simplification was made, with some loss in accuracy, to reduce the operator permutations from 6 to 3, and even down to 2 (*i.e.*, no cross terms associated with vertical transport, as was done in Eq. 4.2 of LR96). This compromise in accuracy undermines a number of the design advantages of the FFSL scheme.

The need to make this 3D generalization is eliminated with the introduction of the Lagrangian control-volume coordinate for the vertical discretization of the 3D hydrostatic flow (Lin and Rood 1998 and Lin and Rood 1999). The Lagrangian control-volume discretization allows us to circumvent many of the problems associated with sigma, pressure, or isentropic coordinates, increasing the physical integrity as well as the computational efficiency of the model. This vertical coordinate is discussed more fully in the next section. (See also D2 Next-Generation Data Assimilation System for structure of the grid, etc.).

With this development the foundation of the next-generation dynamical core is considered mature. Possible shortcomings in dynamical core formulation include the restriction of the problem to hydrostatic flows and the impact of the nonlinear diffusion associated with the monotonicity constraint on long-term climate simulations. Given today's limited computing power, the hydrostatic assumption will not be a limitation for global modeling or data assimilation in the near future.

The impact of the non-linear diffusion associated with the monotonicity constraint is a more difficult problem to embrace. All discrete schemes must address the problem of sub-scale mixing. The finite-volume dynamical core contains a non-linear diffusion that mixes strongly when monotonicity principles are violated. In most numerical schemes, an explicit linear diffusion is added to the equations to provide the subgrid-scale mixing as well as to smooth and/or stabilize the time marching. Numerous other strategies have been used, but the reconciliation of some sort of numerical diffusion with physical processes remains illusive. One manifestation of the diffusion is in the conservation of total energy, which is especially important in the long integrations needed to assess climate change. A method is needed to diagnose and possibly correct any lack of total energy conservation. Such methods are more straightforward in linear schemes. We have experimented with several approaches, including an *ad hoc* global fixer to the total energy after each dynamical time step. In the following section, for the vertical discretization, we present a physically motivated mapping (from the floating Lagrangian to the fixed Eulerian finite control-volume) algorithm that not only conserves the total mass, but also the total energy and momentum.

Before the Lagrangian control-volume concept is introduced in section 3, we summarize here the unique attributes of the finite-volume dynamical core:

- Terrain-following “floating” Lagrangian control-volume vertical coordinate with a monotonicity-preserving and mass-, momentum-, and total energy-conserving mapping algorithm to the fixed Eulerian reference coordinate.

- Genuinely two-dimensional, physically based, conservative semi-Lagrangian transport between two bounding Lagrangian surfaces that define the finite control-volume.

- Accurate finite-volume representation of the mean terrain. Accurate and physically consistent integration of pressure gradient force for the terrain-following Lagrangian control-volume.

3. The floating Lagrangian control-volume vertical coordinate and the mass, momentum, and total energy conserving mapping algorithm

Lagrangian coordinate surfaces are non-penetrable material surfaces. Starr (1945) is perhaps the first to formulate, in the differential form, the governing equations using the terrain-following Lagrangian coordinate. Starr did not apply the Lagrangian control-volume concept nor did he present a solution to the problem of computing the pressure gradient forces. In finite-volume form, the Lagrangian surfaces are the bounding surfaces of the Lagrangian control-volumes within which the finite-volume algorithms developed in LR96, LR97, and L97 can be readily applied to construct the 3D finite-volume dynamical core.

A well-known Lagrangian coordinate is the isentropic coordinate in which the potential temperature (or equivalently, the entropy) is used as the vertical coordinate. Advantages of the isentropic coordinate are documented in the literature. However, there are some inherent disadvantages as well as limitations. First, the isentropic coordinate is not suitable for an unstable, neutral, or nearly neutrally stratified flow (*e.g.*, daytime PBL over desert areas). Second, the pure isentropic coordinate is not terrain following, which calls for the use of massless layers (Hsu and Arakawa 1990) or a complicated hybrid approach with an essentially sigma coordinate near the lower boundary. Third, pure isentropic systems appear to require higher "vertical" resolution in order to resolve the horizontal surface temperature variation by the coordinate surfaces intersecting the ground (Williamson 1998, personal communication). Finally, the diabatic vertical

transport is significant, particularly in the tropical troposphere, which greatly diminishes the primary advantage of the isentropic coordinate system.

A truly Lagrangian coordinate would not have the above-mentioned scientific difficulties. Also, a truly Lagrangian coordinate can reduce the dimensionality of the problem from three to two and facilitate the parallelization of the algorithm. This has far reaching consequences for distributed computing platforms. Realizing that the earth's surface, for all practical modeling purposes, can be treated as a non-penetrable material surface, it becomes straightforward to construct a terrain following Lagrangian control-volume coordinate system. In fact, any commonly used terrain following coordinates (pure sigma, hybrid sigma-p, or hybrid sigma-isentropic coordinate) can be used as the *starting* reference (*i.e.*, fixed) Eulerian coordinate of the floating Lagrangian coordinate system. Figures 3 and 4 on Pages 12 and 13 of the next part of the ATBD, D2 Next-Generation Data Assimilation System, illustrate the actual structure of the Lagrangian vertical coordinate.

The basic idea is to start the time marching from a chosen terrain-following reference Eulerian coordinate (could be any vertical coordinate used in today's GCMs) and to allow the finite-volumes bounded by two coordinate surfaces, *i.e.*, the Lagrangian control-volumes, to float vertically with the flow as dictated by the hydrostatic dynamics. Although we need not use the governing differential equations for the vertical discretization, it is insightful to consider that for hydrostatic flows the "right hand side" and the "left hand side" of the vertical momentum equation should both vanish. That is, if physical consistency is to be exactly maintained, the hydrostatic relation, adopting standard notations

$$\frac{1}{\rho}\frac{\partial p}{\partial z} + g = 0 \qquad (1)$$

requires that

$$\frac{dw}{dt} = 0 \qquad (2)$$

The above two equations indicate that a vertical coordinate that maintains exactly the hydrostatic balance (Eq. 1) would automatically be an inertial Lagrangian coordinate (implied by Eq. 2), and vice versa.

A Lagrangian control-volume is defined as the region bounded from below and above by two neighboring Lagrangian (material) surfaces. From the discrete form of Eq. 1, the pressure thickness δp of that control-volume is proportional to the total mass, i.e.,

$$\delta p = - \rho \, \delta z \tag{3}$$

where the symbol δ represents the difference between the two bounding Lagrangian surfaces. A finite Lagrangian control-volume discretization in the vertical direction, therefore, automatically maintains an exact hydrostatic balance in its discrete form.

The Earth's surface (land and ocean) is the lowest Lagrangian surface and to close the coordinate system the model top, a prescribed constant pressure surface, is assumed to be a Lagrangian surface. As there is no vertical transport across any Lagrangian surface, even when the flow is not adiabatic, a discrete two-dimensional horizontal representation of the basic physical laws (mass conservation for air and water substances, momentum conservation, and the energy conservation, which is the first law of thermodynamics) can then be written for a layer bounded by twoLagrangian surfaces. With the exception of the pressure gradient terms, the finite-volume discretization of the momentum conservation law follows exactly that of the shallow water system (Eq. 14 - Eq. 24 in LR97). The computation of the pressure gradient forces for the Lagrangian control-volume follows exactly L97 (Eq. 11 and Eq. 12 in L97).

In the current formulation of the hydrostatic dynamical core, the potential temperature is chosen as the primary thermodynamic variable for the Lagrangian control-volume. [Alternatively, to ensure energy conservation when the nonlinear diffusion associated with the monotonic transport algorithm is involved, one may consider using the total energy itself as the prognostic thermodynamic variable.] The depth of fluid h in the shallow water system of Lin and Rood (1997) is replaced by the pressure thickness δp. The ideal gas law, the mass conservation law for air mass (δp), the conservation law for the potential temperature, together with the modified momentum equations close the 2D dynamical system, which are vertically coupled via the hydrostatic relation (Eq. 12 in L97).

Given the prescribed pressure at the model top P_∞, the position of each Lagrangian surface P_n (horizontal subscripts omitted) is determined in term of the hydrostatic pressure as follows.

$$P_n = P_\infty + \sum_{k=1}^{n} \delta P_k \ , \qquad (for\ n = 1, 2, 3, \ \ ..., N) \tag{4}$$

where the subscript n is the vertical index ranging from 1 at the lower bounding Lagrangian surface of the first (the highest) layer to N at the Earth's surface. There are N+1 Lagrangian surfaces to define a total number of N Lagrangian layers. The surface pressure, which is the pressure at the lowest Lagrangian surface, is easily computed as P_N using Eq. 4. The surface pressure is needed for the physical parameterizations and to define the reference Eulerian coordinate for the mapping procedure.

As the time proceeds, the coordinate surfaces eventually deform, particularly in the presence of persistent diabatic heating/cooling to such a degree that it will negatively impact the accuracy of the horizontal-to-Lagrangian-coordinate transport and the computation of the pressure gradient forces. The time scale for the deformation is on the order of a few hours to a day. Therefore, a key to the success of the floating Lagrangian coordinate is an accurate and conservative algorithm for mapping the deformed Lagrangian coordinate back to a fixed reference Eulerian coordinate.

There are a number of degrees of freedom in the design of the vertical mapping algorithm. To ensure conservation, the current mapping algorithm is based on the reconstruction of the mass (pressure thickness δp), zonal and meridional winds, tracer mixing ratios, and total energy (volume integrated sum of the internal, potential, and kinetic energy), using a monotonic piecewise parabolic sub-grid distribution with the hydrostatic pressure as the mapping coordinate. Since the mapping is done by integration with the hydrostatic pressure (*i.e.*, mass) as the coordinate, the mapping algorithm is conservative for the quantities to be reconstructed. An outline of the mapping procedure follows.

Step-1: Define a suitable Eulerian reference coordinate. The mass in each layer (δp) is then redistributed according to the chosen Eulerian coordinate. The hybrid sigma-pressure "eta" coordinate currently used in the NCAR CCM3 (Kiehl, et al. 1996) is adopted in the current model setup. (See, Figure 4 on Page13 of the next part of the ATBD, D2 Next-Generation Data Assimilation System.)

Step-2: Construct the vertical subgrid profiles of tracer mixing ratios (Q), zonal and meridional winds (u and v), and total energy (Γ) in the Lagrangian control-volume

coordinate based on the Piece-wise Parabolic Method (PPM, Colella and Woodward 1984).

The total energy Γ is computed as the sum of the finite-volume mean geopotential ϕ, internal energy, and the kinetic energy as follows.

$$\Gamma = \frac{1}{\delta p} \int [C_v T + \phi + \frac{1}{2}(u^2 + v^2)]\, dp$$

Applying integration by parts and the ideal gas law, the above can be rewritten as

$$\Gamma = C_p \overline{T} + \frac{1}{\delta p} \delta (p\phi) + K \qquad (5)$$

where $\overline{T}$ is the layer mean temperature, K is the layer mean kinetic energy, p is the pressure at layer edges, and C_v and C_p are the specific heat of the air at constant volume and at constant pressure, respectively.

Layer mean values of [Q, (u, v), and Γ] in the Eulerian coordinate system are obtained by integrating analytically the sub-grid distributions from model top to the surface, layer by layer.

Step-3: Compute kinetic energy in the Eulerian coordinate system for each layer. Substituting kinetic energy and the hydrostatic relationship (Eq. 13 in L97) into Eq. 5, the layer mean temperature in the Eulerian coordinate is then retrieved from the reconstructed total energy (done in Step-2) by a fully explicit integration procedure starting from the surface up to the model top.

The physical implication of retrieving the layer mean temperature from the total energy described in Step-3 is that the dissipated kinetic energy, if any, is locally converted into internal energy via the vertical sub-grid mixing (dissipation) processes. Due to the monotonicity preserving nature of the sub-grid reconstruction the column-integrated kinetic energy inevitably decreases (dissipates), which then leads to local frictional heating. The frictional heating is a physical process that maintains the conservation of the total energy in a closed system.

As viewed by an observer sitting on the reference Eulerian coordinate, the mapping procedure essentially performs the physical function of the relative-to-the-Eulerian-

coordinate vertical transport, by vertically redistributing mass, momentum, and total energy from the Lagrangian control-volume back to the Eulerian framework.

The model integration cycle consists of one large time step for vertical mapping and several small time steps for the 2D horizontal Lagrangian dynamics. For computational efficiency, the 2D tracer transport is done using the large (mapping) time step. This splitting of the time steps for the dynamics and the tracer transport mimics the process of off-line transport. For transport consistency and mass conservation, the time averaged 2D winds on the C grid and the accumulated horizontal mass fluxes at cell walls are used to perform the large-time-step "off-line" tracer transport. Direct comparisons have been made with "in-line" small-time-step transport, and no significant difference was found even after weeks of integration.

In tests using the Held-Suarez forcing, a three-hour mapping time interval is found to be adequate. In the full model with NCAR CCM3 physics, we used a 30-minute time step for compatibility with the LSM and the physical parameterizations. The size of the small time step for the Lagrangian dynamics depends on the horizontal resolution. For the 2-degree horizontal resolution, a time step size of 450 seconds can be safely used for the Lagrangian dynamics.

From the large-time-step transport standpoint, the small-time-step integration of the 2D Lagrangian dynamics may be regarded as an accurate iterative solver for computing the time mean winds and the mass fluxes, analogous in functionality to a semi-implicit algorithm's elliptic solver. The merit of "explicit" versus "semi-implicit" algorithm ultimately comes down to the relative efficiency of each approach. In light of the advantage in parallelization, we do not consider our fully explicit algorithm for the Lagrangian dynamics as a liability in computational efficiency. Furthermore, it may be possible to further increase the size of the small time step via vertical mode decomposition. This approach is one of the algorithm design issues we plan to revisit.

Compared with DAO's operational model in the GEOS-3 system, a center difference algorithm constructed on the Arakawa C-grid, the finite-volume dynamical core based model does not suffer from the computational instability near poles that so profoundly degraded the quality of the stratospheric analysis that an expensive polar rotation algorithm was implemented to maintain the quality of the analysis. Furthermore, there is no need for any type of filtering for the removal of two-grid-interval waves. Due to the

Lagrangian vertical structure and the horizontal FFSL algorithm, a two to three times larger time step (as compared to GEOS-3 GCM) can be used for the 2D Lagrangian dynamics. The larger dynamical time step and the computational advantages associated with the vertical Lagrangian discretization enable the new model to have a 5 to 10 times better throughput than the current GEOS-3 GCM (comparisons performed on the SGI origin-2000 shared memory machine).

4. Physical parameterizations

The NCAR CCM3 physical parameterizations represent a well-balanced set of processes with a long history of development and documentation (Kiehl et al., 1996 for CCM3, Hack et al., 1993 for CCM2 and references therein for additional details and characteristics of individual components). The scientific algorithms and the characteristics of the climate simulations of the NCAR CCM3 and the coupled atmosphere-ocean-land model, the "Climate System Model", have also been well documented in a special issue of the *Journal of Climate* (Vol. 11, No, 6, 1998). Efforts are underway to develop a new "Common Land Model", a revised cloud overlap scheme, a turbulence kinetic energy based atmospheric boundary layer parameterization, and a revised gravity wave drag scheme suitable not only for tropospheric climate simulations but also for stratospheric and mesospheric applications.

In the current setup of the joint model the dynamics and all physical parameterizations are operator-split as follows:

$$q^{n+1} = W(T(S(R(M(D(q^{n})))))) \qquad (6)$$

where q represents the prognostic variables (the state vector), D the finite-volume dynamical core, M the moist physics packages (cumulus convection with both updrafts and downdrafts, mid and shallow convection, and the large-scale condensation and rain re-evaporation), R the cloud and radiation packages (diagnostic clouds; short and long-wave radiation parameterizations), S the surface models (prescribed sea-ice and SST; the land surface model), T the PBL mixing/turbulence parameterization, and W is the gravity wave drag parameterization. The model is effectively two-time-level with each operator

(from *D* to *W* in alphabetic order) performing instant adjustments to the state vector *q*. The proposed *rapid update cycle* for the interface to the analysis segment of the data assimilation system will fit naturally into the above scheme as the last operator (or equivalently the first). (See, page 18 of the next part of the ATBD, D2 Next-Generation Data Assimilation System.) This update strategy is designed to minimize interference between various physical components and the analysis which would simplify the tuning and the development processes for the data assimilation system.

5. Software development

Some of the organizational issues of software development are discussed in Appendices C-E of Part A of the ATBD, Overview). This section focuses on issues of parallelization.

Extensive efforts in the parallelization of the model software are currently underway at the DAO, Lawrence Livermore National Laboratory (LLNL), and NCAR. DAO is leading the scientific as well as the software developments of the finite-volume dynamical core, which has been successfully parallelized using the state-of-the-art hybrid Message Passing Interface (MPI) and the OpenMP approach. This hybrid approach is adopted in light of the trend in the US computing hardware industry toward the clustering of shared memory machines built on commodity microprocessors. The current MPI implementation is applied only to the meridional direction for a 1D decomposition. The OpenMP multi-tasking (or multi-threading) is applied mostly in the vertical direction. The model can be run with MPI alone, OpenMP alone, or hybrid MPI and OpenMP. Due to the vertical Lagrangian control-volume discretization, on the SGI Origin-2000, the parallel efficiency using purely OpenMP is in fact better than using purely MPI. However, the best parallel efficiency is obtained with the combination (*i.e.*, the hybrid) of MPI and OpenMP when large number of processors are to be used.

To get an idea of the current shared memory multitasking (OpenMP) capability, Figure 1 shows the relative percentage of wall-clock-time taken by the dynamical core (yellow bars), the I/O (purple bars), and the CCM3 physical parameterizations (red bars) in the 2^o x 2.5^o (latitude x longitude) 32-layer joint model setup with various total number of CPUs. The time step used was 30 min. Due to better scaling, the percentage of time taken by the finite-volume dynamical core decreases with increasing CPU number

(from near 50% using single CPU to under 40% with 32 CPUs). With 32 CPUs, the full model's speed up is about 20 (*vs*. ~24 with dynamical core only setup), and the throughput is about two years per wall-clock-day on the 250 MHz SGI Origin-2000. At 1^o x 1.25^o and 32-level, the throughput using 32 CPUs is about 120 days per day, which already exceeds DAO's current operational requirement.

Work in implementing the Land Surface Model into the hybrid MPI-OpenMP framework is still underway. It is estimated that the full model will be able to efficiently use 256 to 512 CPUs with the hybrid approach on cluster SMP machines (*e.g*., the SGI Origin or the IBM SP3) with throughput roughly an order of magnitude better than the ones given above using OpenMP only parallelization.

The limitation with only an explicit 1D decomposition using MPI is that the number of usable processors is quite limited on a pure distributed memory computing platform (*e.g*., the CRAY T3E). DAO is therefore collaborating with LLNL on development of a 2D (meridional and vertical) MPI decomposition. The goal is for the model at the 1-degree horizontal resolution and 55 layers to scale to ~1000 CPUs on both pure distributed memory computing platform and cluster shared memory platform.

6. Model results

The initial research of the model jointly developed with NCAR focuses on the impact on climate simulations of different dynamical cores (currently three) interfaced with the same physics packages from the NCAR CCM3. Several 18-level, 2^o x 2.5^o simulations were performed to provide direct comparison with the T42, 18-level (model top at 2.9 mb) standard CCM3 configuration. In addition, similar experiments were performed with an experimental semi-Lagrangian CCM3. Currently all direct comparisons with the NCAR Eulerian and/or semi-Lagrangian dynamics were run for 5 years with climatological sea surface temperature (SST). Further comparisons with standard NCAR CCM will be made at a higher vertical resolution, using different cumulus parameterizations, and with multi-year analyzed SST. Output data files from the joint model have been sent to NCAR, and similar files from CCM3 have been provided to NASA. Model evaluation is proceeding at both institutions.

For application in DAO's troposphere-stratosphere data assimilation system, we extended NCAR CCM3's standard 18-level setup to a 55-level "middle atmosphere" configuration with model top at 0.01 mb. Long term climate simulations using the 2^o x 2.5^o and 1^o x 1.25^o horizontal resolutions have been carried out. Basic mean climates produced at the two different resolutions were found to be very similar except at polar latitudes, where the typical model cold bias is substantially reduced with the 1^o x 1.25^o resolution. In addition, the 1^o x 1.25^o climate run also produced realistic tropical cyclones (animations of intense tropical cyclones generated by the 1^o x 1.25^o NASA/NCAR model can be found at http://dao.gsfc.nasa.gov/NASanCAR/). The 2^o x 2.5^o L55 configuration was chosen as the baseline model for the participation in the AMIP2 comparison (http://www-pcmdi.llnl.gov/amip/). We show exemplary results below to establish that the model is performing credibly with no overt pathologies. More extensive diagnostics to establish the model's strengths and weaknesses are a high priority in the development effort.

The initial focus of the model validation was on the basic climate of the troposphere. Figure 2 compares the 15-year mean DJF (December-January-February) 500 mb eddy height (defined as mean height with zonal means removed) from the AMIP2 run using the 55-layer joint NASA/NCAR model (the upper panel), and the 15-year ECMWF reanalysis (the lower panel). Without extensive re-tuning of the NCAR CCM physics, the model simulation captures very well the basic structure seen in ECMWF reanalysis. The ability to simulate this pattern was of special concern to NCAR scientists as prototype versions of CCM4, using the same physics package, were much too zonal.

Another concern with the CCM is the weak tropical variability and the lack of a clear MJO (Madden-Julian Oscillation) signal (see, *Journal of Climate*, Vol. 11, No, 6, 1998). Figure 3 shows the time-longitude diagram of the band-pass filtered 200 mb velocity potential in the tropics from the AMIP2 run. Although the propagation speed is slightly faster than the observed, there is a clear and strong MJO signal with magnitude as strong as any models participating in the AMIP comparison (Slingo, 1996). The MJO is a phenomenon involving nonlinear interaction of the dynamics and the parameterized physical processes. Given previous experiments focusing on the use of spectral dynamics and a set of cumulus parameterizations, the success of the MJO simulation may be

viewed as being as sensitive to the interaction of the physics and the dynamics as it is the specifics of the convection algorithms.

In order for both NCAR and DAO to have interest in continued model development, it was not adequate to just establish that the joint model is capable of credible climate simulations. It is also necessary to assure that the model does not suffer any fatal flaw in numerical weather prediction applications. Therefore, we have also carried out a few 5-day forecast experiments using the joint model at the 2^o x 2.5^o L55 resolution. The initial conditions were obtained by interpolating the 2^o x 2.5^o L70 analyses produced by the GEOS-2 DAS without any additional initialization procedure. As such, the objective was not to get the best possible forecast. Rather, the goal was to detect, if any, major flaws in the scientific algorithms or the software of the joint model. Figure 4 compares the 500 mb geopotential height anomaly correlation scores from the 2^o x 2.5^o L55 joint model, the 2^o x 2.5^o L70 GEOS-2 GCM, and a mid 90s version of NCEP's operational medium range forecast model. The scores for the northern and the southern extra-tropics are shown in the upper panel and the lower panel, respectively. The joint model produces a better forecast than the other two models for this test case, which was arbitrarily chosen. The results from two other cases, also arbitrarily chosen, were similarly good. Given the limited experiments, the conclusion we can draw at this point is that the joint model with new generation dynamical core and the NCAR CCM3 physics is capable of both credible weather predictions and climate simulation. Given that the NCAR physics were not designed for NWP applications and no tuning of the joint model for forecast applications, the quality of these forecasts are very encouraging.

Many of the DAO's customers use assimilated data products for chemistry-transport applications. Scientists at NCAR also have a special interest in constituent problems. Both tropospheric and stratospheric applications are of interest. We examine below some basic characteristics of the model simulation that have direct implication to the stratospheric chemistry and transport simulations. This example also explicitly reveals some basic changes that have followed from the new dynamical core.

Figure 5 compares the simulated temperature (red curve) in January at 100 mb with the ECMWF reanalysis (ERA-15, the black curve) and 13 other climate models in the GRIPS (GCM-reality Inter-comparison Project for SPARC) project (green curves; data from the GRIPS models provided by S. Pawson 1999, personal communication). The

joint model with the 55-level setup is performing very well with practically no bias away from the poles. Near the summer pole the joint model has a cold bias, which is smaller than most of the GRIPS-participating models. Similar comparisons have also been performed for different time periods, and the quality of the simulation from the joint model is equally good.

The transport of tracers from the troposphere to the stratosphere is an important problem for both chemistry and climate applications. The simulation of the “atmospheric tape recorder” as pictured by (Mote et al., 1996) has become a central diagnostic of this transport. This focus arises from the study of, for instance, Hall et al. (1999) who show that, for all models evaluated, the atmospheric residence time of long-lived tracers is too short and tracers propagate upwards too rapidly. Past experiences in 3D chemistry transport using GEOS-1 and GEOS-2 DAS produced winds are consistent with this general pathology. CH_4 and N_2O transport, for instance, across the tropical tropopause is too strong and too fast as compared to the observed.

Figure 6 compares the model simulated "atmospheric tape recorder" (Mote et al. 1996) in the joint model (labeled as FVCCM3) and the GEOS-2 GCM. Both model runs were for the same AMIP2 period with the same boundary forcing (prescribed multi-year SST and sea-ice) and the same horizontal resolution. Shown is the water vapor anomaly, difference from long-term average. Water vapor is effectively conservative in the tropical lower stratosphere. The assent rate of the "tape" (the water vapor anomaly) in the joint model (FVCCM3) is very near that indicated by the data (Mote et al. 1996). On the other hand, the assent rate in the GEOS-2 GCM simulated "tape recorder" is about 30% to 50% faster than that indicated by the data.

To investigate this further, Figure 7 compares the zonal means of the vertical velocity in the same vertical and horizontal domain for both models. Both models exhibit a similar annual cycle in the lower stratosphere. Interestingly, even though the joint model has a slower upward transport of water vapor, the overall upward velocity is stronger. The assent rate of the water vapor signal is, in fact, consistent with the vertical velocity in the joint model. Also interesting is the annual cycle of the vertical velocity in GEOS-2 shows a large negative (downward) phase between 100 mb and 70 mb during the mid-year (the northern summer), which is likely unrealistic.

Figure 8 compares the simulated monthly mean vertical velocity for January 1980 at 100 mb for both models. The horizontal structure of the vertical velocity from the GEOS-2 run is dominated by noise over a large portion of the globe. Some small-scale noises are also present in the joint model, but the extent is limited to the mountainous regions where orographically generated upward propagating gravity waves may be expected. Given the overall consistency of the transport with the velocity field in the joint model, and a concomitant inconsistency in GEOS-2, we assert that the small-scale noise is causing spurious transport that moves tracers into the stratosphere too rapidly. We believe that this better treatment of the vertical velocity is directly related to the formulation of the next-generation dynamical core, and in particular, those design features that reduce spurious pressure gradient forces near regions of high topography.

The comparisons shown in this section are representative of many initial evaluations of the performance of the joint model, and we feel safe to conclude that the joint model with the finite-volume dynamics and the Lagrangian vertical structure provides a credible simulation even without extensive reworking of the NCAR CCM3 physics parameterizations. Additional results from the joint model as well as data are accessible from the internet (http://dao.gsfc.nasa.gov/NASanCAR/).

7. Summary

We have presented the formulation and initial results from the first version of the prototype next-generation model being developed jointly with the National Center for Atmospheric Research (NCAR). The dynamical formulation is unique in global atmospheric studies and is based on a finite-volume approach to computational fluid dynamics. The finite volumes are bound by piecewise continuous functions which allows analytic integration around volumes to represent conservation laws. The vertical structure of the model relies on a floating Lagrangian coordinate that provides a high degree of physical consistency as well as numerous computational advantages.

Results have been presented from climate, forecast, and tracer applications. Results from each of these applications are promising, and offer strategies to address known shortcomings of existing models at both NCAR and DAO. Presently, a prototype effort to identify any fatal weaknesses in data assimilation applications is underway. This prototype is described in the next section of the ATBD, D2 Next-Generation Data

Assimilation System. Initial results from the prototype are expected by the time of the panel review.

References:

Allen, D. J., A. R. Douglass, R. B. Rood, and P. D. Guthrie, 1991: Application of a monotonic upstream-biased transport scheme to three-dimensional constituent transport calculations. *Mon. Wea. Rev.*, **119**, 2456-2464.

Allen, D. J., P. J. Kasibhatla, A. M. Thompson, R. B. Rood, B. G. Doddridge, K. E. Pickering, R. D. Hudson, and S.-J. Lin, 1996: Transport induced interannual variability of carbon monoxide determined using a chemistry and transport model. *J. Geophys. Res.*, **101**, 28655-28669.

Bonan, G. B., 1996: A land surface model for ecological, hydrological, and atmospheric studies: technical description and user's guide. *NCAR Technical Note NCAR/TN-417+STR.* National Center for Atmospheric Research, Boulder, Colorado, 150 pp.

Colella, P., and P. R. Woodward, 1984: The piecewise parabolic method (PPM) for gas-dynamical simulations. *J. Comput. Phys.*, **54**, 174-201.

Douglass, A. R., R. B. Rood, S. R. Kawa, and D. J. Allen, 1997: A three-dimensional simulation of the evolution of the middle latitude winter ozone in the middle atmosphere. *J. Geophys. Res.,* **102,** 19217-19232.

Hack, J. J., B. A. Boville, J. T. Kiehl, P. J Rasch and D. L. Williamson, 1993: Description of the NCAR Community Climate Model (CCM2).NCAR Technical Note, NCAR/TN-382+STR, Boulder, CO, 108 pp.

Hall, T. M., D. W. Waugh, K. A. Boering, and R. A. Plumb, 1999: Evaluation of transport in stratospheric models, *J. Geophys. Res.*, **104,** 18,815-18,839.

Hsu, Y.-J. G., and A. Arakawa, 1990: Numerical modeling of the atmosphere with an isentropic vertical coordinate. *Mon. Wea. Rev.*, **118**, 1933-1959.

Kiehl, J. T., J. J. Hack, G. B. Bonan, B. A. Boville, B. P. Briegleb, D. L. Williamson, and P. J. Rasch 1996: Description of the NCAR Community Climate Model (CCM3). NCAR Technical Note, NCAR/TN-420+STR, Boulder, CO, 152pp.

Lin, S.-J., W. C. Chao, Y. C. Sud, and G. K. Walker, 1994: A class of the van Leer-type transport schemes and its applications to the moisture transport in a general circulation model. *Mon. Wea. Rev*., **122**, 1575-1593.

Lin, S.-J., and R. B. Rood, 1996: Multidimensional Flux Form Semi-Lagrangian Transport schemes. *Mon. Wea. Rev.*, 124, 2046-2070.

Lin, S.-J., 1997: A finite-volume integration method for computing pressure gradient forces in general vertical coordinates. *Q. J. Roy. Met. Soc*., **123**, 1749-1762.

Lin, S.-J., and R. B. Rood, 1997: An explicit flux-form semi-Lagrangian shallow water model on the sphere. *Q. J. Roy. Met. Soc.*, **123**, 2477-2498.

Lin, S.-J., and R. B. Rood, 1998: A flux-form semi-Lagrangian general circulation model with a Lagrangian control-volume vertical coordinate. The Rossby-100 symposium, Stockholm, Sweden.

Lin, S.-J., and R. B. Rood, 1999: Development of the joint NASA/NCAR General Circulation Mode. Preprint, 13th conference on Numerical Weather Prediction, Denver, CO.

Machenhauer, B., 1998: MPI workshop on conservative transport schemes. Max-planck-Institute for Meteorology. Report No. 265.

Mote, P. W., and coauthors, 1996: An atmospheric tape recorder: The imprint of tropical tropopause temperatures on stratospheric water vapor. *J. Geophys. Res*., **101**, No. D2, 3989--4006

Rood, R. B., 1987: Numerical advection algorithms and their role in atmospheric transport and chemistry models. *Rev. Geophys*., **25**, 71-100.

Rood, R. B., and Coauthors, 1991: Three-dimensional simulations of wintertime ozone variability in the lower stratosphere. *J. Geophys. Res*., **96**, 5055-5071.

Slingo, J. M., and Coauthors, 1996: Intraseasonal oscillations in 15 atmospheric general circulation models: results from an AMIP diagnostic subproject. *Climate Dynamics*, **12**, 325-357.

Staniforth, A., and J. Cote, 1991: Semi-Lagrangian integration schemes for atmospheric models ---A review. *Mon. Wea. Rev*., **119**, 2206-2223.

Starr, V. P., 1945: A quasi-Lagrangian system of hydrodynamical equations. *J. Meteor*., **2**, 227-237.

Van Leer, B., 1977: Toward the ultimate conservative difference scheme. Part IV: A new approach to numerical convection. *J. Comput. Phys*., **23**, 276-299.

Williamson, D. L., J. G. Olson, and B. A. Boville, 1998: A comparison of semi-Lagrangian and Eulerian tropical climate simulations. *Mon. Wea. Rev*., vol **126**, 1001—1012.

Woodward, P. R., and P. Colella, 1984: The numerical simulation of two-dimensional fluid flow with strong shocks. *J. Comput. Phys*., **54**, 115-173.

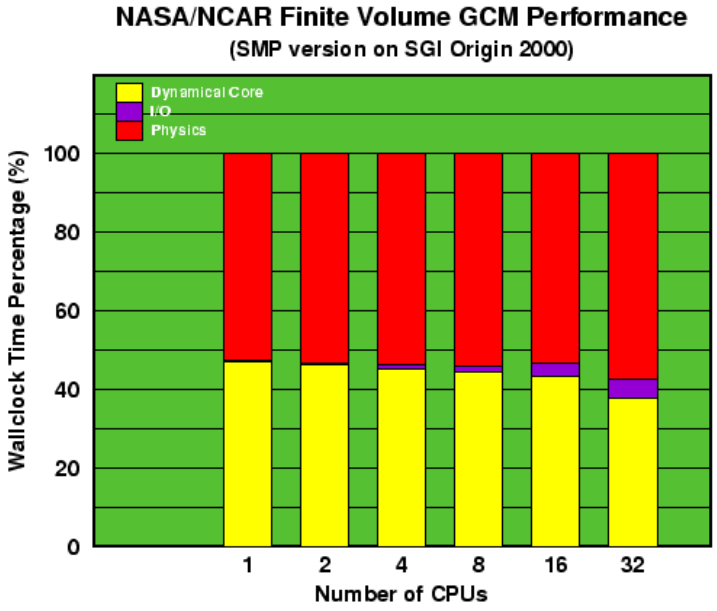

Figure 1. Breakdown of the total wall-clock-time between the finite-volume dynamical core (yellow bars), the I/O (purple bars), and the CCM3 physical parameterizations (red bars) in the 32-layer configuration of the joint NASA/NCAR model.

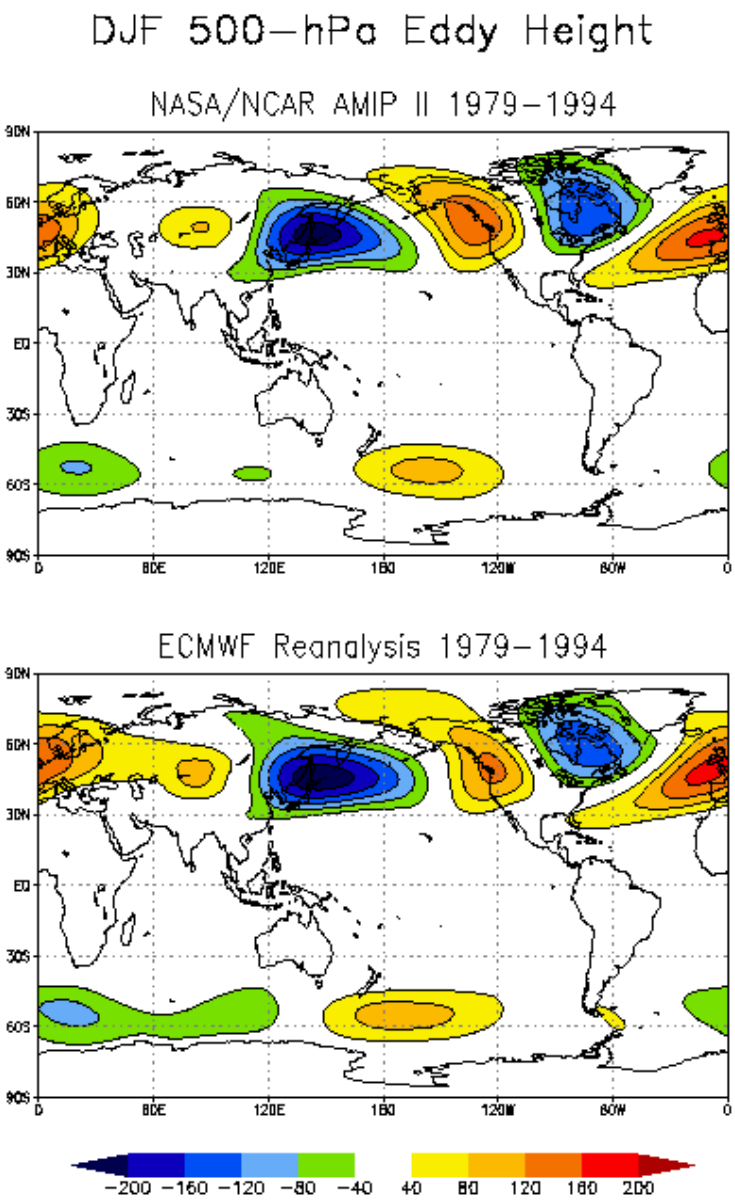

Figure 2. DJF 500 mb eddy height fields from the AMIP2 run using the 55-layer joint model (top panel), and the ECMWF 15-year reanalysis (bottom panel).

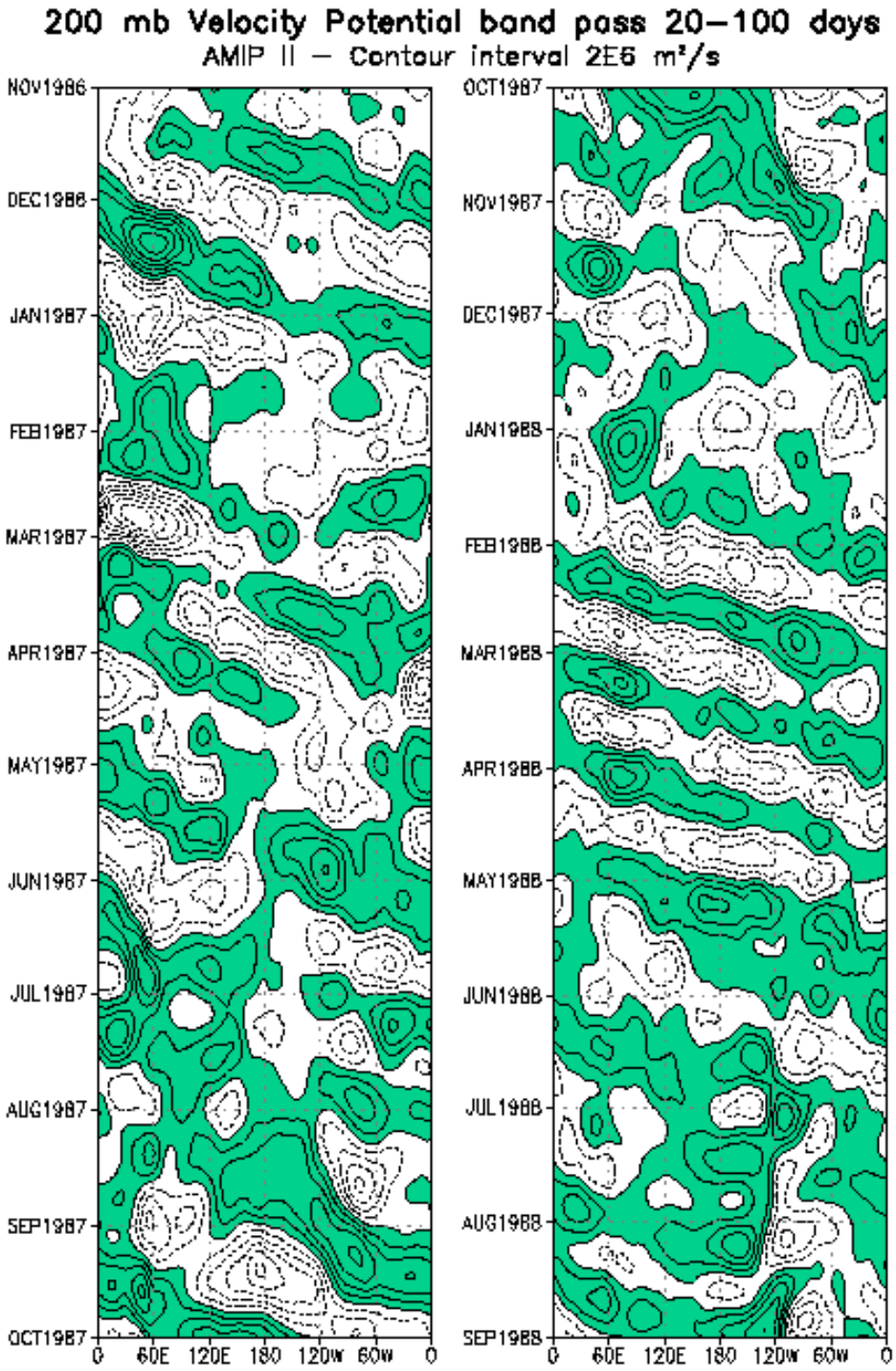

Figure 3. Time-longitude diagram of the band-pass filtered 200 mb velocity potential in the tropics (from 6S to 6N) from the AMIP2 run using the 55-layer joint model.

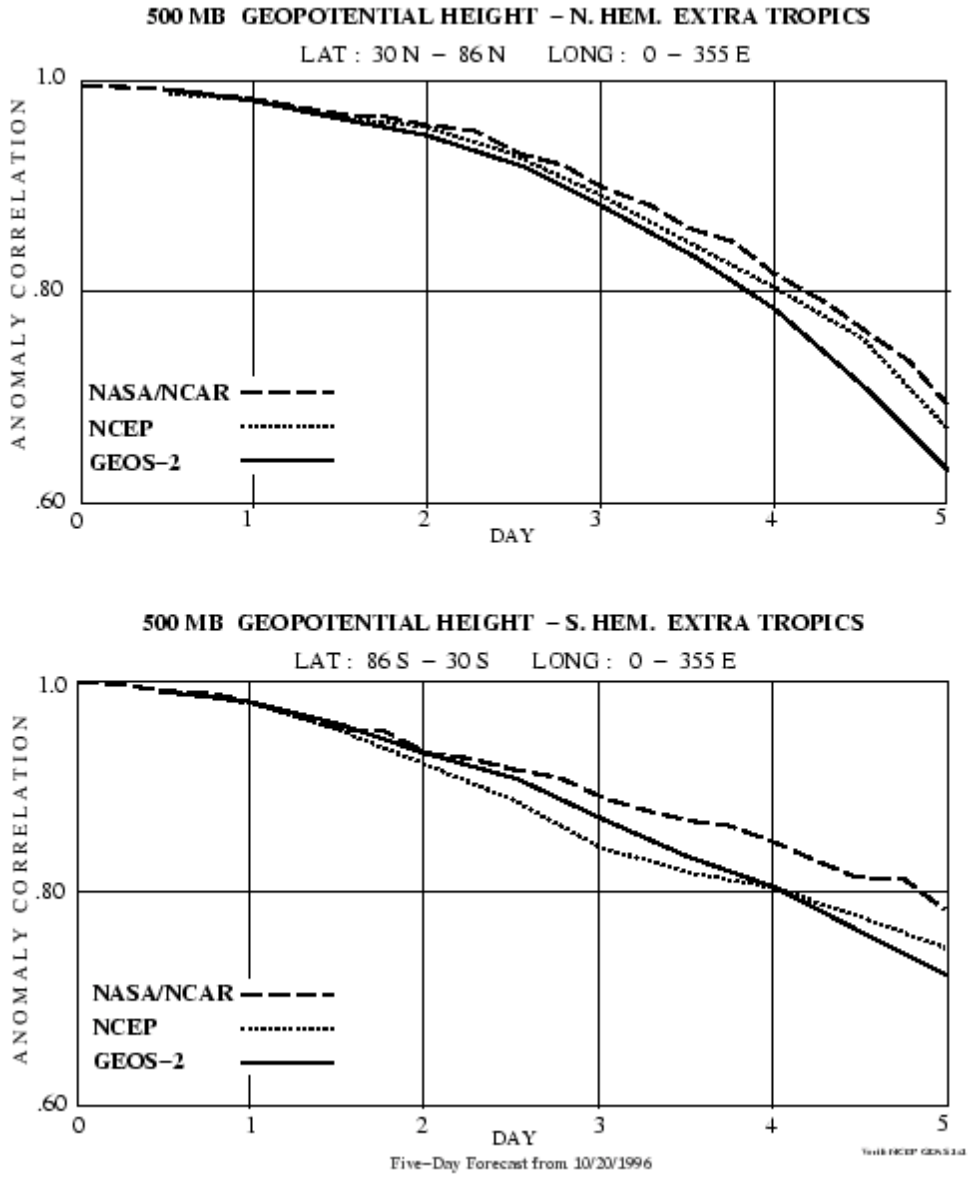

Figure 4. Anomaly correlation scores for the 500 mb geopotential height for a 5-day period starting from Oct. 20, 1996. The upper and the lower panels are for the northern extra tropics and the southern extra tropics, respectively. Forecasts from three models are shown: the joint NASA/NCAR model (long dashed), the NCEP model (short dashed), and the GEOS-2 GCM (solid).

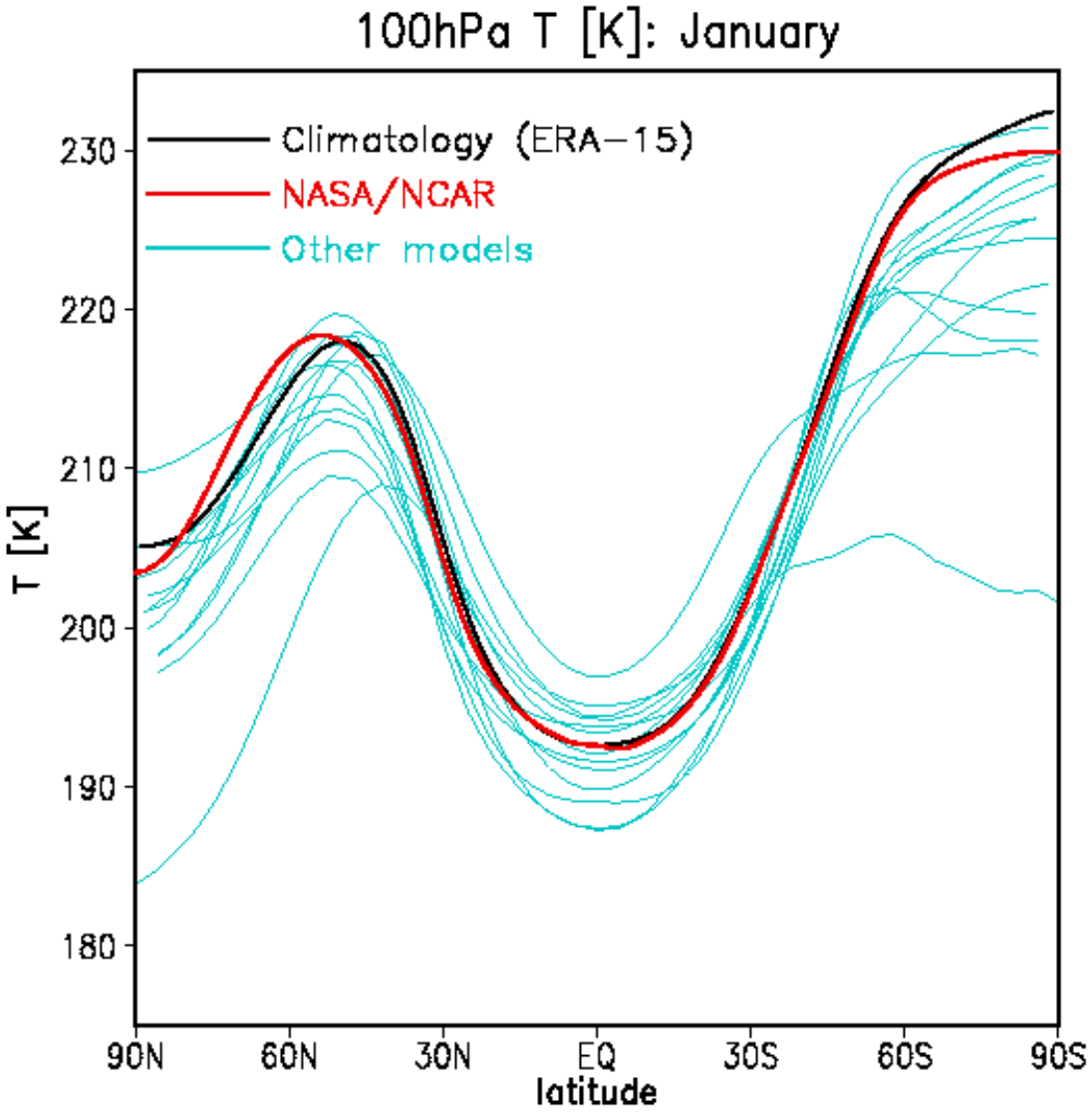

Figure 5. Comparisons of the simulated January 100 mb temperature from the 55-layer version of the joint model (red curve), 13 other GRIP models (green curves), and the ECMWF 15-year reanalysis (black curve).

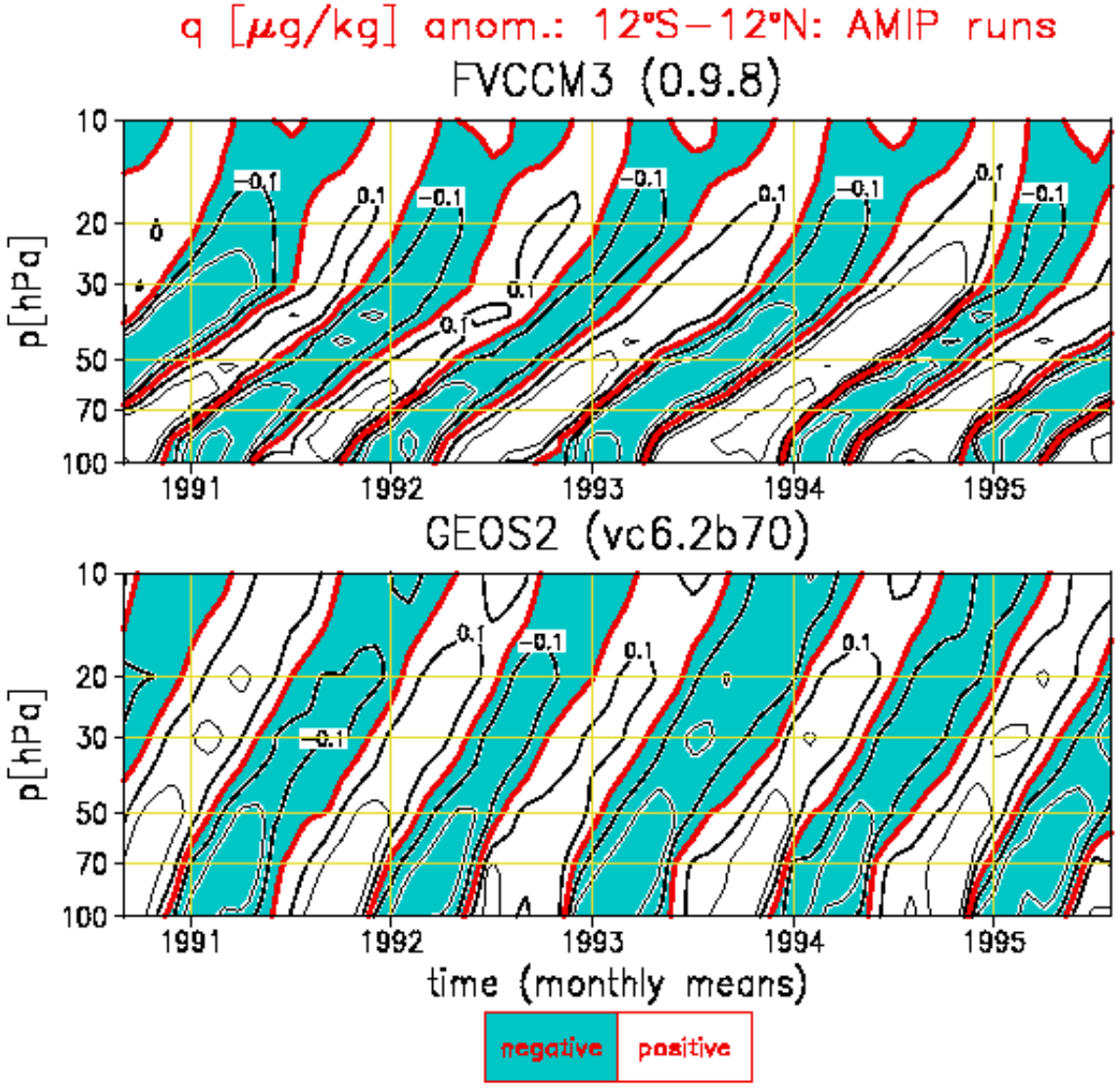

Figure 6. The time-pressure sections of the equatorial water vapor anomaly in the lower stratosphere (The "tape recorder", Mote et al. 1996) as simulated by the joint NASA/NCAR model (the upper panel), and the GEOS-2 GCM (the lower panel).

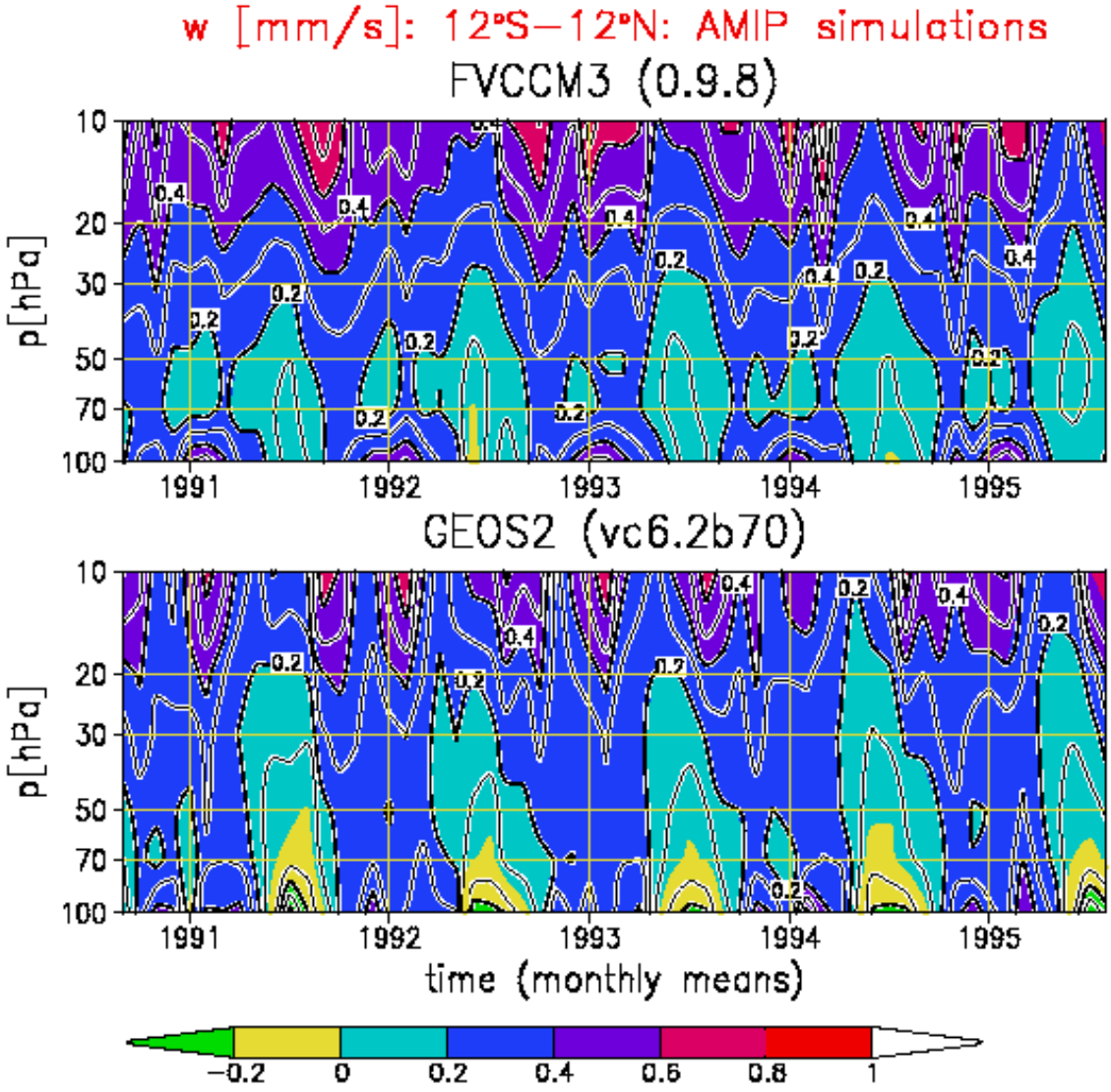

Figure 7. As in Figure 5, but for the zonal mean vertical velocity (converted to mm/s from mb/s using a constant scale height of 7 km).

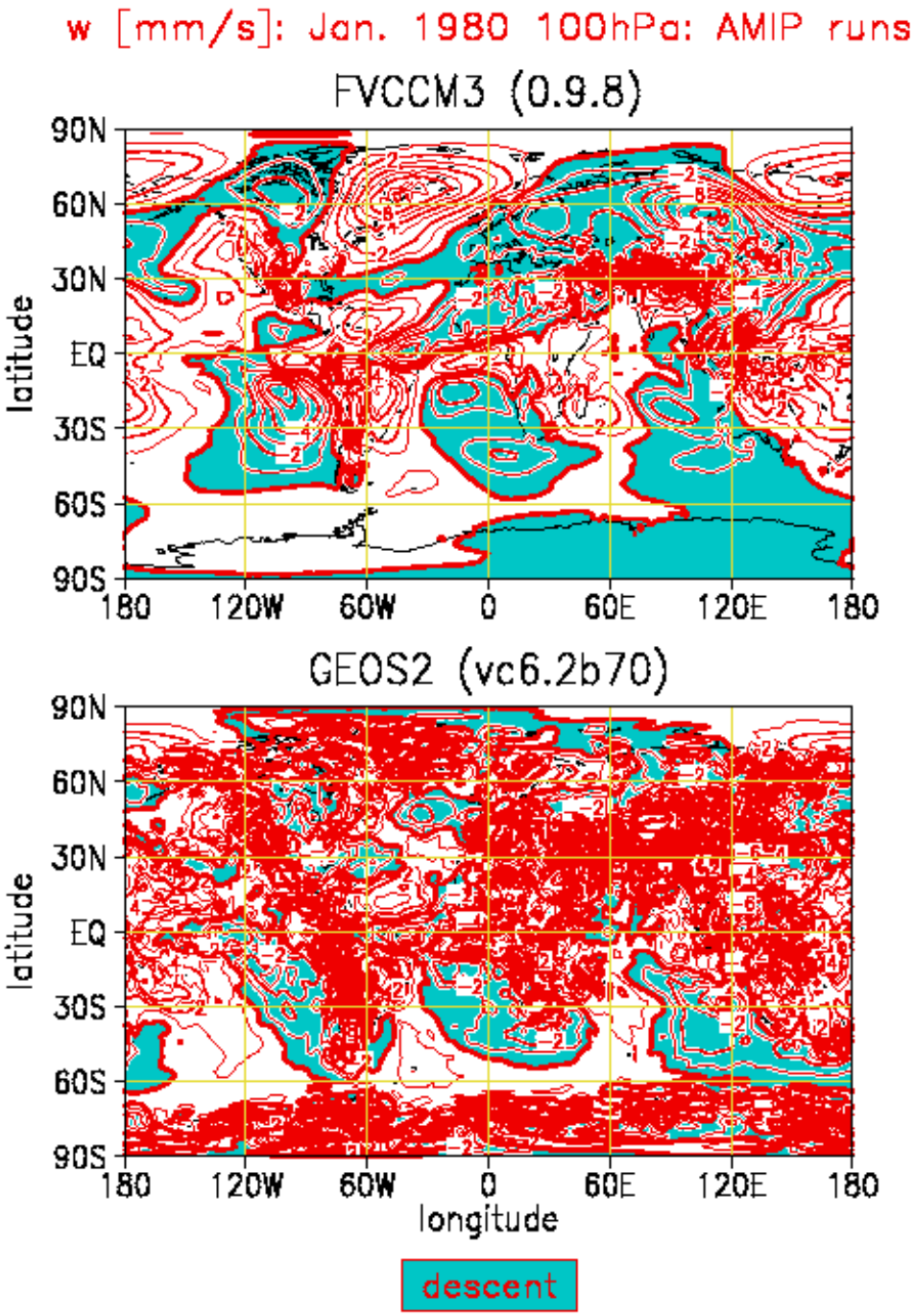

Figure 8. The monthly mean vertical velocity (converted to mm/s from mb/s using a constant scale height of 7 km) for January 1980 at 100 mb as simulated by the joint NASA/NCAR model (upper panel, labeled as FVCCM3), and the GEOS-2 GCM (lower panel).